

# SEQUENCE LISTING

<110> PILETZ, John E.  
IVANOV, Tina R.

<120> DNA MOLECULES ENCODING IMIDALINE RECEPTIVE POLYPEPTIDES  
AND POLYPEPTIDES ENCODED THEREBY

<130> Corrected Sequence Listing

<140> 08/922,635  
<141> 1997-09-03

<150> 08/650,766  
<151> 1996-05-20

<150> 60/012,600  
<151> 1996-03-01

<160> 22

<170> PatentIn Ver. 2.0

<210> 1  
<211> 3385  
<212> DNA  
<213> Homo sapiens  
<220>  
<221> CDS  
<222> (1398)..(3383)

<400> 1

```
gctctagaac tagtg gatcc cccgggctgc aggaattcca gtttaatact aacccta atg 60
tgtgactgcg gtttacaaag agctctgtat cacctgggat agctttcagt agcaattcac 120
tacaactggt cctaaaaaat aataacaata ataataataa ttagagaatt aaaacccaac 180
agcatgttga atggttaaaa tcacgtaaga actgaaattt ggggtggggg tgcctcaac 240
agctgagctt gtcctagcag tgaaaatgct cgctccaag cagggtcag aaaggtctgg 300
agccctccag gcagagggt gagctcagg ggctcttga ggacactcac cccatggtcc 360
```



gcc Ala	agc Ser	agc Ser 110	ctc Leu	tcg Ser	tcc Ser	act Thr	gac Asp 115	agt Ser	ctg Leu	act Thr	ccc Pro	gag Glu 120	cac His	cag Gln	ccc Pro	1766
att Ile	gcc Ala 125	cag Gln	gga Gly	tgt Cys	tct Ser	gat Asp 130	tcc Ser	ttg Leu	gag Glu	tcc Ser	atc Ile 135	cct Pro	gcg Ala	gga Gly	cag Gln	1814
gca Ala 140	gct Ala	tcc Ser	gat Asp	gat Asp 145	tta Leu	agg Arg	gac Asp	gtg Val	cca Pro	gga Gly 150	gct Ala	gtt Val	ggg Gly	ggg Gly	gca Ala 155	1862
agc Ser	cca Pro	gaa Glu	cat His	gcc Ala 160	gag Glu	ccg Pro	gag Glu	gtc Val	cag Gln 165	gtg Val	gtg Val	ccg Pro	ggg Gly	tct Ser 170	ggc Gly	1910
cag Gln	atc Ile	atc Ile 175	ttc Phe	ctg Leu	ccc Pro	ttc Phe	acc Thr	tgc Cys 180	att Ile	ggc Gly	tac Tyr	acg Thr 185	gcc Ala	acc Thr	aat Asn	1958
cag Gln	gac Asp 190	ttc Phe	atc Ile	cag Gln	cgc Arg	ctg Leu	agc Ser 195	aca Thr	ctg Leu	atc Ile	cgg Arg 200	cag Gln	gcc Ala	atc Ile	gag Glu	2006
cgg Arg	cag Gln 205	ctg Leu	cct Pro	gcc Ala	tgg Trp	atc Ile 210	gag Glu	gct Ala	gcc Ala	aac Asn 215	cag Gln	cgg Arg	gag Glu	gag Glu	ggc Gly	2054
cag Gln 220	ggg Gly	gaa Glu	cag Gln	ggc Gly	gag Glu 225	gag Glu	gag Glu	gat Asp	gag Glu	gag Glu 230	gag Glu	gaa Glu	gaa Glu	gag Glu	gag Glu 235	2102
gac Asp	gtg Val	gct Ala	gag Glu	aac Asn 240	cgc Arg	tac Tyr	ttt Phe	gaa Glu	atg Met 245	ggg Gly	ccc Pro	cca Pro	gac Asp	gtg Val 250	gag Glu	2150
gag Glu	gag Glu	gag Glu	gga Gly 255	gga Gly	ggc Gly	cag Gln	ggg Gly	gag Glu 260	gaa Glu	gag Glu	gag Glu	gag Glu	gaa Glu 265	gag Glu	gag Glu	2198
gat Asp	gaa Glu 270	gag Glu	gcc Ala	gag Glu	gag Glu	gag Glu	cgc Arg 275	ctg Leu	gct Ala	ctg Leu	gaa Glu 280	tgg Trp	gcc Ala	ctg Leu	ggc Gly	2246
gcg Ala	gac Asp 285	gag Glu	gac Asp	ttc Phe	ctg Leu	ctg Leu 290	gag Glu	cac His	atc Ile	cgc Arg 295	atc Ile	ctc Leu	aag Lys	gtg Val	ctg Leu	2294
tgg Trp 300	tgc Cys	ttc Phe	ctg Leu	atc Ile	cat His 305	gtg Val	cag Gln	ggc Gly	agt Ser	atc Ile 310	cgc Arg	cag Gln	ttc Phe	gcc Ala	gcc Ala 315	2342
tgc Cys	ctt Leu	gtg Val	ctc Leu	acc Thr 320	gac Asp	ttc Phe	ggc Gly	atc Ile	gca Ala 325	gtc Val	ttc Phe	gag Glu	atc Ile	ccg Pro 330	cac His	2390
cag Gln	gag Glu	tct Ser 335	cgg Arg	ggc Gly	agc Ser	agc Ser	cag Gln 340	cac His	atc Ile	ctc Leu	tcc Ser	tcc Ser	ctg Leu 345	cgc Arg	ttt Phe	2438

gtc Val	ttt Phe	tgc Cys 350	ttc Phe	ccg Pro	cat His	ggc Gly	gac Asp 355	ctc Leu	acc Thr	gag Glu	ttt Phe	ggc Gly 360	ttc Phe	ctc Leu	atg Met	2486
ccg Pro	gag Glu 365	ctg Leu	tgt Cys	ctg Leu	gtg Val	ctc Leu 370	aag Lys	gta Val	cgg Arg	cac His	agt Ser 375	gag Glu	aac Asn	acg Thr	ctc Leu	2534
ttc Phe 380	att Ile	atc Ile	tcg Ser	gac Asp	gcc Ala 385	gcc Ala	aac Asn	ctg Leu	cac His	gag Glu 390	ttc Phe	cac His	gcg Ala	gac Asp	ctg Leu 395	2582
cgc Arg	tca Ser	tgc Cys	ttt Phe	gca Ala 400	ccc Pro	cag Gln	cac His	atg Met	gcc Ala 405	atg Met	ctg Leu	tgt Cys	agc Ser	ccc Pro 410	atc Ile	2630
ctc Leu	tac Tyr	ggc Gly	agc Ser 415	cac His	acc Thr	agc Ser	ctg Leu	cag Gln 420	gag Glu	ttc Phe	ctg Leu	cgc Arg	cag Gln 425	ctg Leu	ctc Leu	2678
acc Thr	ttc Phe	tac Tyr 430	aag Lys	gtg Val	gct Ala	ggc Gly	ggc Gly 435	tgc Cys	cag Gln	gag Glu	cgc Arg	agc Ser 440	cag Gln	ggc Gly	tgc Cys	2726
ttc Phe 445	ccc Pro	gtc Val	tac Tyr	ctg Leu	gtc Val	tac Tyr 450	agt Ser	gac Asp	aag Lys	cgc Arg	atg Met 455	gtg Val	cag Gln	acg Thr	gcc Ala	2774
gcc Ala 460	ggg Gly	gac Asp	tac Tyr	tca Ser	ggc Gly 465	aac Asn	atc Ile	gag Glu	tgg Trp	gcc Ala 470	agc Ser	tgc Cys	aca Thr	ctc Leu	tgt Cys 475	2822
tca Ser	gcc Ala	gtg Val	cgg Arg	cgc Arg 480	tcc Ser	tgc Cys	tgc Cys	gcg Ala	ccc Pro 485	tct Ser	gag Glu	gcc Ala	gtc Val	aag Lys 490	tcc Ser	2870
gcc Ala	gcc Ala	atc Ile	ccc Pro 495	tac Tyr	tgg Trp	ctg Leu	ttg Leu	ctc Leu 500	acg Thr	ccc Pro	cag Gln	cac His	ctc Leu 505	aac Asn	gtc Val	2918
atc Ile	aag Lys	gcc Ala 510	gac Asp	ttc Phe	aac Asn	ccc Pro	atg Met 515	ccc Pro	aac Asn	cgt Arg	ggc Gly	acc Thr 520	cac His	aac Asn	tgt Cys	2966
cgc Arg	aac Asn 525	cgc Arg	aac Asn	agc Ser	ttc Phe	aag Lys 530	ctc Leu	agc Ser	cgt Arg	gtg Val	ccg Pro 535	ctc Leu	tcc Ser	acc Thr	gtg Val	3014
ctg Leu 540	ctg Leu	gac Asp	ccc Pro	aca Thr	cgc Arg 545	agc Ser	tgt Cys	acc Thr	cag Gln	cct Pro 550	cgg Arg	ggc Gly	gcc Ala	ttt Phe	gct Ala 555	3062
gat Asp	ggc Gly	cac His	gtg Val	cta Leu 560	gag Glu	ctg Leu	ctc Leu	gtg Val	ggg Gly 565	tac Tyr	cgc Arg	ttt Phe	gtc Val	act Thr 570	gcc Ala	3110
atc Ile	ttc Phe	gtg Val	ctg Leu 575	ccc Pro	cac His	gag Glu	aag Lys	ttc Phe 580	cac His	ttc Phe	ctg Leu	cgc Arg	gtc Val 585	tac Tyr	aac Asn	3158

cag	ctg	cgg	gcc	tcg	ctg	cag	gac	ctg	aag	act	gtg	gtc	atc	gcc	aag	3206
Gln	Leu	Arg	Ala	Ser	Leu	Gln	Asp	Leu	Lys	Thr	Val	Val	Ile	Ala	Lys	
		590					595					600				
acc	ccc	ggg	acg	gga	ggc	agc	ccc	cag	ggc	tcc	ttt	gcg	gat	ggc	cag	3254
Thr	Pro	Gly	Thr	Gly	Gly	Ser	Pro	Gln	Gly	Ser	Phe	Ala	Asp	Gly	Gln	
	605					610					615					
cct	gcc	gag	cgc	agg	gcc	agc	aat	gac	cag	cgt	ccc	cag	gag	gtc	cca	3302
Pro	Ala	Glu	Arg	Arg	Ala	Ser	Asn	Asp	Gln	Arg	Pro	Gln	Glu	Val	Pro	
	620				625					630					635	
gca	gag	gct	ctg	gcc	ccg	gcc	cca	gtg	gaa	gtc	cca	gct	cca	gcc	ccg	3350
Ala	Glu	Ala	Leu	Ala	Pro	Ala	Pro	Val	Glu	Val	Pro	Ala	Pro	Ala	Pro	
				640					645						650	
gaa	ttc	gat	atc	aag	ctt	atc	gat	acc	gtc	gac	ct					3385
Glu	Phe	Asp	Ile	Lys	Leu	Ile	Asp	Thr	Val	Asp						
			655					660								

<210> 2  
 <211> 1954  
 <212> DNA  
 <213> Homo sapiens

<400> 2

atgacaggcc aggtgggagc gcagacggtc tcgggtggca aaagaagcat tgcaggtctg 60  
 acacttgatga ggccgctcag aagtgtgcac ctgcttgata tgtccgttca agtgatcagg 120  
 cctggagaag cctttccac agctctggca gatgtaaggc ggaattcccc agagaagaag 180  
 ggtggtgaag actcccggct ctgagctgcc ccctgcatca gaccagcag ctcccctccc 240  
 actgtggctc ccgcatctgc ctccctgccc cagcccatcc tctctaacca aggaatcatg 300  
 ttcggttcagg aggaggccct ggccagcagc ctctcgtcca ctgacagtct gactcccag 360  
 caccagccca ttgcccaggg atgtttctgat tccttgagat ccatccctgc gggacaggca 420  
 gcttccgatg atttaaggga cgtgccagga gctgttggtg gtgcaagccc agaacatgcc 480  
 gagccggagg tccaggtggt gccgggggtct ggccagatca tcttctgccc ctccacctgc 540  
 attggctaca cggccaccaa tcaggacttc atccagcgcc tgagcacact gatccggcag 600  
 gccatcgagc ggcagctgcc tgcttgatc gaggtgccca accagcggga ggagggccag 660  
 ggtgaacagg gcgaggagga ggatgaggag gaggaagaag aggaggacgt ggctgagaac 720  
 cgctactttg aaatggggcc ccagacgtg gaggaggagg agggaggagg ccagggggag 780  
 gaagaggagg aggaagagga ggatgaagag gccgaggagg agcgctggc tctggaatgg 840  
 gccctgggag cggacgagga ctctctgctg gagcacatcc gcatcctcaa ggtgctgtgg 900









ggctgcttcc ccgtctacct ggtctacagt gacaagcgca tggcgcagac ggccgcccggg 600  
 gactactcag gcaacatcga gtggggccagc tgcacactct gttcagccgt gcggcgctcc 660  
 tgctgcgcgc cctctgaggc cgtcaagtcc gccgccatcc cctactggct gttgctcacg 720  
 cccagcacc tcaacgtcat caaggccgac ttcaacccca tgcccaaccg tggcaccac 780  
 aactgtcgca accgcaacag cttcaagctc agccgtgtgc cgctctccac cgtgctgctg 840  
 gacccacac gcagctgtac ccagcctcgg ggcgcctttg ctgatggcca cgtgctagag 900  
 ctgctcgtgg ggtaccgctt tgtcactgcc atcttcgtgc tgccccacga gaagttccac 960  
 ttctgcgcgc tctacaacca gctgcggggc tcgctgcagg acctgaagac tgtggtcac 1020  
 gccaagaccc ccgggacggg aggcagcccc cagggtcctt ttgcggatgg ccagcctgcc 1080  
 gagcgcaggg ccagcaatga ccagcgtccc caggaggtcc cagcagaggc tctggccccg 1140  
 gccccagtgg aagtcccagc tccagccccg g 1171

<210> 5  
 <211> 651  
 <212> PRT  
 <213> Homo sapiens

<400> 5

Met	Thr	Gly	Gln	Val	Gly	Ala	Gln	Thr	Val	Ser	Gly	Gly	Lys	Arg	Ser	1	5	10	15
Ile	Ala	Gly	Leu	Thr	Leu	Val	Arg	Pro	Leu	Arg	Ser	Val	His	Leu	Leu	20	25	30	
Asp	Met	Ser	Val	Gln	Val	Ile	Arg	Pro	Gly	Glu	Ala	Phe	Pro	Thr	Ala	35	40	45	
Leu	Ala	Asp	Val	Arg	Arg	Asn	Ser	Pro	Glu	Lys	Lys	Gly	Gly	Glu	Asp	50	55	60	
Ser	Arg	Leu	Ser	Ala	Ala	Pro	Cys	Ile	Arg	Pro	Ser	Ser	Ser	Pro	Pro	65	70	75	80
Thr	Val	Ala	Pro	Ala	Ser	Ala	Ser	Leu	Pro	Gln	Pro	Ile	Leu	Ser	Asn	85	90	95	
Gln	Gly	Ile	Met	Phe	Val	Gln	Glu	Glu	Ala	Leu	Ala	Ser	Ser	Leu	Ser	100	105	110	
Ser	Thr	Asp	Ser	Leu	Thr	Pro	Glu	His	Gln	Pro	Ile	Ala	Gln	Gly	Cys	115	120	125	
Ser	Asp	Ser	Leu	Glu	Ser	Ile	Pro	Ala	Gly	Gln	Ala	Ala	Ser	Asp	Asp	130	135	140	
Leu	Arg	Asp	Val	Pro	Gly	Ala	Val	Gly	Gly	Ala	Ser	Pro	Glu	His	Ala	145	150	155	160

Glu	Pro	Glu	Val	Gln	Val	Val	Pro	Gly	Ser	Gly	Gln	Ile	Ile	Phe	Leu
				165					170					175	
Pro	Phe	Thr	Cys	Ile	Gly	Tyr	Thr	Ala	Thr	Asn	Gln	Asp	Phe	Ile	Gln
			180					185					190		
Arg	Leu	Ser	Thr	Leu	Ile	Arg	Gln	Ala	Ile	Glu	Arg	Gln	Leu	Pro	Ala
		195					200					205			
Trp	Ile	Glu	Ala	Ala	Asn	Gln	Arg	Glu	Glu	Gly	Gln	Gly	Glu	Gln	Gly
	210					215					220				
Glu	Glu	Glu	Asp	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Asp	Val	Ala	Glu	Asn
225					230					235					240
Arg	Tyr	Phe	Glu	Met	Gly	Pro	Pro	Asp	Val	Glu	Glu	Glu	Glu	Gly	Gly
				245					250					255	
Gly	Gln	Gly	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Asp	Glu	Glu	Ala	Glu
			260					265						270	
Glu	Glu	Arg	Leu	Ala	Leu	Glu	Trp	Ala	Leu	Gly	Ala	Asp	Glu	Asp	Phe
		275					280					285			
Leu	Leu	Glu	His	Ile	Arg	Ile	Leu	Lys	Val	Leu	Trp	Cys	Phe	Leu	Ile
	290					295					300				
His	Val	Gln	Gly	Ser	Ile	Arg	Gln	Phe	Ala	Ala	Cys	Leu	Val	Leu	Thr
305					310					315					320
Asp	Phe	Gly	Ile	Ala	Val	Phe	Glu	Ile	Pro	His	Gln	Glu	Ser	Arg	Gly
				325					330					335	
Ser	Ser	Gln	His	Ile	Leu	Ser	Ser	Leu	Arg	Phe	Val	Phe	Cys	Phe	Pro
			340					345					350		
His	Gly	Asp	Leu	Thr	Glu	Phe	Gly	Phe	Leu	Met	Pro	Glu	Leu	Cys	Leu
		355					360					365			
Val	Leu	Lys	Val	Arg	His	Ser	Glu	Asn	Thr	Leu	Phe	Ile	Ile	Ser	Asp
	370					375					380				
Ala	Ala	Asn	Leu	His	Glu	Phe	His	Ala	Asp	Leu	Arg	Ser	Cys	Phe	Ala
385					390					395					400
Pro	Gln	His	Met	Ala	Met	Leu	Cys	Ser	Pro	Ile	Leu	Tyr	Gly	Ser	His
				405					410					415	
Thr	Ser	Leu	Gln	Glu	Phe	Leu	Arg	Gln	Leu	Leu	Thr	Phe	Tyr	Lys	Val
			420					425					430		
Ala	Gly	Gly	Cys	Gln	Glu	Arg	Ser	Gln	Gly	Cys	Phe	Pro	Val	Tyr	Leu
		435					440					445			
Val	Tyr	Ser	Asp	Lys	Arg	Met	Val	Gln	Thr	Ala	Ala	Gly	Asp	Tyr	Ser
	450					455					460				
Gly	Asn	Ile	Glu	Trp	Ala	Ser	Cys	Thr	Leu	Cys	Ser	Ala	Val	Arg	Arg
465					470					475					480
Ser	Cys	Cys	Ala	Pro	Ser	Glu	Ala	Val	Lys	Ser	Ala	Ala	Ile	Pro	Tyr
				485					490					495	

Trp Leu Leu Leu Thr Pro Gln His Leu Asn Val Ile Lys Ala Asp Phe  
 500 505 510  
 Asn Pro Met Pro Asn Arg Gly Thr His Asn Cys Arg Asn Arg Asn Ser  
 515 520 525  
 Phe Lys Leu Ser Arg Val Pro Leu Ser Thr Val Leu Leu Asp Pro Thr  
 530 535 540  
 Arg Ser Cys Thr Gln Pro Arg Gly Ala Phe Ala Asp Gly His Val Leu  
 545 550 555 560  
 Glu Leu Leu Val Gly Tyr Arg Phe Val Thr Ala Ile Phe Val Leu Pro  
 565 570 575  
 His Glu Lys Phe His Phe Leu Arg Val Tyr Asn Gln Leu Arg Ala Ser  
 580 585 590  
 Leu Gln Asp Leu Lys Thr Val Val Ile Ala Lys Thr Pro Gly Thr Gly  
 595 600 605  
 Gly Ser Pro Gln Gly Ser Phe Ala Asp Gly Gln Pro Ala Glu Arg Arg  
 610 615 620  
 Ala Ser Asn Asp Gln Arg Pro Gln Glu Val Pro Ala Glu Ala Leu Ala  
 625 630 635 640  
 Pro Ala Pro Val Glu Val Pro Ala Pro Ala Pro  
 645 650

<210> 6  
 <211> 390  
 <212> PRT  
 <213> Homo sapiens

<400> 6

Glu Glu Glu Glu Glu Glu Asp Glu Glu Ala Glu Glu Glu Arg Leu Ala  
 1 5 10 15  
 Leu Glu Trp Ala Leu Gly Ala Asp Glu Asp Phe Leu Leu Glu His Ile  
 20 25 30  
 Arg Ile Leu Lys Val Leu Trp Cys Phe Leu Ile His Val Gln Gly Ser  
 35 40 45  
 Ile Arg Gln Phe Ala Ala Cys Leu Val Leu Thr Asp Phe Gly Ile Ala  
 50 55 60  
 Val Phe Glu Ile Pro His Gln Glu Ser Arg Gly Ser Ser Gln His Ile  
 65 70 75 80  
 Leu Ser Ser Leu Arg Phe Val Phe Cys Phe Pro His Gly Asp Leu Thr  
 85 90 95  
 Glu Phe Gly Phe Leu Met Pro Glu Leu Cys Leu Val Leu Lys Val Arg  
 100 105 110

[illegible]



<210> 12  
<211> 20  
<212> DNA  
<213> Homo sapiens

<400> 12

ccattttaaa gtagcggttc

20

<210> 13  
<211> 20  
<212> DNA  
<213> Homo sapiens

<400> 13

aggagagaaa ggtgagccaa

20

<210> 14  
<211> 20  
<212> DNA  
<213> Homo sapiens

<400> 14

gtagatcctg aggttgacca

20

<210> 15  
<211> 20  
<212> DNA  
<213> Homo sapiens

<400> 15

tgtgagcatt tctggccttc

20

<210> 16  
<211> 20  
<212> DNA  
<213> Homo sapiens

<400> 16

tgaagacgcc agagaagcag

20

<210> 17  
<211> 20  
<212> DNA  
<213> Homo sapiens

<400> 17

gcctcacaag tgtcagacct 20

<210> 18  
<211> 18  
<212> DNA  
<213> Homo sapiens

<400> 18

agaaggggtgg tgaagact 18

<210> 19  
<211> 20  
<212> DNA  
<213> Homo sapiens

<400> 19

cttggttaga gaggatgggc 20

<210> 20  
<211> 20  
<212> DNA  
<213> Homo sapiens

<400> 20

gcccatacctc tctaaccaag 20

<210> 21  
<211> 15202  
<212> DNA  
<213> Homo sapiens

<400> 21

gatccgagct caattaaccc tacttaaagg gagtcgactc gataccttaa atattcatat 60  
ctcctggaca acctgtggcc atagtgcctg actgtaaacc caaaggggtt gcctttgcca 120  
gtgtagccca gcctggtgtc tgctgcccct cgcggtgtct gtgcacctgc cacgatgctg 180





aggcccatc gtctgccac tatcttagcg ttttcaaagg gctttcacct ctgaaccag 2100  
gcatcctcgg agatgagtga gtgaagcagg tctcatgagc gtgtctgctg gcccgcccc 2160  
cacggaagag gggaggggtgt gccgtccga gtggagccga ggctcgggac acgcaggaaa 2220  
ggacgccgcc tgcccggtt cctggagacg cagaacttgg tgtgaggtct tgggaaaaca 2280  
gttcaaccgg atgttttaag agccagaaaa acattccac cccttgacct ggtaacccca 2340  
ctggtgggga ttttctctta gagggataag ataccgggaa ggggaggtga aatgctcacc 2400  
actgcaaaaa cacgggctgc aactgcaaca tcggaggatg agaggagag tcggctgtgg 2460  
tgcagaatgc tcagcagccc tccagcagg gacaggaaga ctgggcagga agaggggaga 2520  
agcattcaag ttaaggcaaa aggcccaacg cagagcagca cactgaggtc acacctgtga 2580  
gatgtggaag agaattcctg agcgtggagc gatgggggta ggtgccagga tgattgccca 2640  
ttttgcttct gtcagactct tgactaagga tttctggtg cattttatta cataaaagcc 2700  
agggaggtta tatcacgtg agaaagcttc cctgacgcc cctcctgtag cgcagccaag 2760  
cgagcctgtg gaggtaccat atgactgtag gcctctggg acaggagct gcatctgctt 2820  
ctcaaggcca gggacacagc catttctgcc agcatctgtt gatcagtga tgagtgaagt 2880  
ggcaggtaga gcaggagcca gtgaagagca ggccctggat ggggtgggat gcaccatgtc 2940  
cccaggctgc agctgcaggc agccccccac attgtcggag aagcctctgc accagctcag 3000  
ccccctctc actcccctt tgccctggg acactctgca gaggggcact ctgcagtctg 3060  
tccccgcat cgttgactt ctggacatgg cctccagatt tgcacctctt aaataaatct 3120  
gcagtggatg tctttgtgtg cacctctctt tccttttgg gagaaacagc aaagatcgga 3180  
cccctaagga ctctcctgat gtctcgtct tatccgctga gtgcccttct tgaccacttg 3240  
tttgtacagg ccacggtcca ggacgggagc agatagactg tcctgtccc tgtccacatt 3300  
tccttggtcc aaacagggt tgtgggaggt agtgcaaaa ggtgttggtc tttttctcac 3360  
tgatttgag gcctcccgt gtgttttttc agccgctgt tcctgggtct tgccctggatg 3420  
gacagggttt tttagcgcgt gggagcagct ttgctgacca tgcctgttgc ttccagcctg 3480  
attcccagga agggagcgtg cttgcgaagg aactggcact cgggcctgcc tgaagggggc 3540  
gctgtccaga cacaccagc ctcccgctgt ggcaggcgt gtcggagcca tggatgattg 3600  
tgaccaatag ggggtggtcg cagagttgat tgtccagcca ggcccaggg ctgagaggag 3660  
gctgtgtgga gaggtggtta ggagccagg ctcggtcagc tgagtctgca tgccagcttc 3720  
ctagctgtgg gacctcaagc aactgttagc ccctctgaag ctgttttctc aactgtgaag 3780  
tgagcgacc ctacttcatt gattctaaga ggcacgcat tccacctgt gacttctctg 3840  
aaactgaggt gcgtctttca gtcagtggcg tctcatagtc gctgtcagcc agctggtatt 3900

cgagatggag tcgtggaaaa cccgtggaca ccttccgcta ggaccaagat ggcgccacct 3960  
gccgcatctt agatttgatg aaatgtggta aataacgaga ggcatgcatg agcgaatgct 4020  
ggggaggcgc ttggcactac ccagagctcc acagagggtg tcgatgaggg ctgccctttc 4080  
ccacatcctt agtagggggg tcaacatgac ccagactgtg cccctgggga gcttgagacc 4140  
atgcgggagg atgagccatg tgctggagga gaacagggtg ggatggtgtg gggcttttgt 4200  
agactgtcta gagcagagaa ggtctgcagt ggagggtgtg tctgaggtga atctcgaagg 4260  
tgaataggag ttgaacgtta gcaggcagag ggtggattgc aggagagcag cggcctgggc 4320  
aggtgcccag cgtggcccat cagggtgctt catgcatggc tgtgtgcttg ccatccttcc 4380  
tgctgccta cccctgctg ctctgcttca tgggggcgtt tgagcttggg cccacctgcc 4440  
tgctcgctt gtgggcagag gaccaggt gtgtgagttg tcctgtcccg gggagcagct 4500  
gagcttgtcc gggttcctcg acctgtgggg cttcagagga cttcgggtca tttcaatggg 4560  
ctgtggcgat gctggctgtg gaggtagcct agggctcctg tagccttcag tgagactggc 4620  
ggcccgatgc ccagtgttca ccctgctggc ggcagtcagg aacatgttca caaagcttta 4680  
cttcaagtgg tctagaggtg atctgaggtg gagtaacagg tccagatagg ctacgttcat 4740  
aaaacagctt cagcgggggt taggaacact gtgcatttac gggacgcagt gggtcagagt 4800  
gctgctgtcc gtgggaggtg gcccagggc aggtcagtg gcacgtcctg tggtaagtgg 4860  
gactgtggat gtgggctcag gctggactca gcagccctgc tggataccaa ggcctgcaag 4920  
ggctggcccc ctggtgaatt gtcccgtgcc ctgtgtatct atgagtcctg cagagatgac 4980  
aaatcagggg acgggggtcat gtctagtcac cgtctgggaa aatgctccag gagtgaacac 5040  
atttcaggct cttgatggat gtacctocaa actcttctct ggatgggtgg gccagcttgc 5100  
atgcctgtgc cggcctctgc ccagcagagt cagggccagg ccacacagtc agtctgactt 5160  
tggcagaagt tgagaggcaa cacttgtctc ttgtttcagc ttgcctttct ttgtgtactt 5220  
ctgagagcga gcattctttt catgtttctat ccgctggccg ttcttctgcg gaatgtctgt 5280  
tcacgtcctt tgcagtctgt taatgaggtt tccaaccttc cctcattttt gtaatctgta 5340  
agaacttttt ccagactagc gatataaata cttgtcaaata attgcaaaca cttttctcat 5400  
ttcatctggt tttaatctat cctgggtttt aaaaaatgtg tctgtggaag tttaattttt 5460  
atgtagtcac atctcagttt ttttccattg catttattct cagaatgctt ctccctgccc 5520  
tgagattaga taagcagtca tttgttcttt cttgagttat tttgagattt cagttttaac 5580  
attttcttct ataatccatg tggctgggtt ttgggatctg gctaaccccc gccatgccag 5640  
tagcctgagg ggcccagccc cacttggtga acagccgctc tccccgcccc acccaccctg 5700  
cctgcctgcc caccgcct ggtctctcca ggaatcatgt tcgttcagga ggaggccctg 5760







taccagaccc cagggctgga ggtgactgcc tgcgtgctgc tctccaccaa ggctgtgtac 11400  
 tttgtgctcc acgacggcct ccgccgctac ttctcagagc cactgcaggg taggcacagg 11460  
 gcctgctggg gctcaggagc ttggagtgtg tgggtggggc aggctgggg ggctattctc 11520  
 tggagccagc tgtgtggctt caggcagcag tcagcgactt ggctgcagtg ggctgagagt 11580  
 tccttgtctg aggaaggag ctgtcatgag ggaggggtcc atggccagat gtgaacgcag 11640  
 aatgcactga gccagggcct ggtgactgct tgggaacagc ctgtgatgag aaggggttag 11700  
 gcagcctttg ccctggggc tgcacaggaa gccctagcca gcgacctgg gactccctg 11760  
 agctggaaga ggctcagact ccagagggca ttgcctatgg ggctttgcac ggggtggaagc 11820  
 caggccagcc aagaggacct gttcctgctg gatgtgctgc acacctagga accttgtgct 11880  
 tgctgccac cgctccctc tgctcccttc tccatcacac agatttctgg catcagaaaa 11940  
 acaccgacta caacaacagc cctttccaca tctcccagtg cttcgtgcta aagcttagtg 12000  
 acctgcagtc agtcaatgtg gggcttttcg accagcattt ccggctgacg cgtgggtgac 12060  
 cctctgtgct ttgtcctatt tcgggtgaag gccagcatca ccagtgggct tccaccttc 12120  
 gtacgtgggt gggttatcat agacagttat ctctgtgctc aagagccact tcttaccgg 12180  
 ggtgggagga agcagcttca ggaactgctg agagagcaga actcacgctc cagggtcag 12240  
 agcaggaggt aggggtgtgcg gcaagcgctg gcccggacag aagcagagtg ggccctggtc 12300  
 tcgggcagga tgtttctgac tcacatttcc tgaggagaga aagctaagct ctttgcctaa 12360  
 tgtctctgtc tccccttcca gaaaaatgcc tcagctcttc cggcctgaag gaatggcctc 12420  
 ctcccgggcc ccatgattct ttctgtgtg ggcctcctg gccctggcct ctgggctgag 12480  
 gcttgctagg gactcggggg ggctctaagg ggcagggata gggctgggga gcgccggcct 12540  
 gtggccctga ccagccctt ctctgtcagg ttccacccc atgcagggtg tcacgtgctt 12600  
 gacgcgggac agctacctga cgcactgctt cctccagcac ctcatggctg tgctgtcctc 12660  
 tctggaacgc acgccctgc cggagcctgt tgacaaggac ttctactccg agtttgggaa 12720  
 caagaccaca ggtaccctg tctagctcag gctgcagaca ggctgcctgg acagacgtca 12780  
 tgggccccag ggtggctctc tgtgccccag aaccctctct gcctctatgt ctctcttttc 12840  
 tcacttagct ggccaggggt ttatgtgggg cttttcgatg gcagagtctc cactccagca 12900  
 gtccctcaac catctggcag acacatctcc agtgctgct ttgggctcct ggctgtggg 12960  
 cccacactt ggagcatcct ctctgctg tctcatgccg ggtctctcg gttggcttg 13020  
 ggcccttggg gctcccagcc ccaccaggg cgggttccag gctatagccc aggtggcatc 13080  
 tctctgcagg gaagatggag aactacgagc tgatccactc tagtcgcgtc aagtttacct 13140  
 accccagtga ggaggagatt ggggacctga cgttcactgt ggccaaaag atggctgagc 13200

cagagaaggc	cccagccctc	agcatccctgc	tgtacgtgca	ggccttcacg	gtgggcatgc	13260
caccccctgg	gtgctgcagg	ggccccctgc	gccccaaagac	actcctgctc	accagctccg	13320
agatcttcct	cctggatgag	gactgtgtcc	actaccact	gcccagagttt	gccaaagagc	13380
cgccgcagag	agacagggtac	cggtctggacg	atggccgcgcg	cgtccggggac	ctggaccgag	13440
tgctcatggg	ctaccagacc	taccgcgagg	ccctcacct	cgtcttcgat	gacgtgcaag	13500
gtcatgacct	catgggcagt	gtcaccctgg	accactttgg	ggaggtgcc	ggtggcccgg	13560
ctagagccag	ccagggccgt	gaagtccagt	ggcaggtgtt	tgtccccagt	gctgagagca	13620
gagagaagct	catctcgctg	ttggctcgcc	agtgggaggc	cctgtgtggc	cgtgagctgc	13680
ctgtcgagct	caccggctag	cccaggccac	agccagcctg	tcgtgtccag	cctgacgcct	13740
actggggcag	ggcagcaggc	ttttgtgttc	tctaaaaatg	ttttatcctc	cctttggtac	13800
cttaatttga	ctgtcctcgc	agagaatgtg	aacatgtgtg	tgtgttgtgt	taattctttc	13860
tcattgttggg	agtgagaatg	ccgggcccct	cagggtgtgc	ggtgtgtgt	cagcctccca	13920
caggtggtac	agccgtgcac	accagtgtcg	tgtctgctgt	tgtgggaccg	ttgttaacac	13980
gtgacactgt	gggtctgact	ttctcttcta	cacgtccttt	cctgaagtgt	cgagtccagt	14040
cctttgttgc	tgttgctgtt	gctgttgcgt	ttgctgttgg	catcttgcgt	ctaatacctga	14100
ggctggtagc	agaatgcaca	ttggaagctc	ccaccccata	ttgttcttca	aagtggaggt	14160
ctcccctgat	ccagacaagt	gggagagccc	gtgggggcag	gggacctgga	gctgccagca	14220
ccaagcgtga	ttcctgctgc	ctgtattctc	tattccaata	aagcagagtt	tgacaccgtc	14280
tgcattcttct	aaaccaaggg	tcaactgggat	cgagtcgacg	gccctatagt	gagtcgtatt	14340
agagctcgcg	gccgcgagct	ctagatgcat	gctcgagcgg	ccgccagtgt	gatggatatc	14400
tgcagaattc	cagcacactg	gcggccgtta	ctagtggatc	cgagctccac	agaggtggtc	14460
gatgagggct	gccctttccc	acatcccttag	taggggggttc	aagatgaccc	agactgtgcc	14520
cctggggagc	ttggagccat	gcgggaggat	gagccatgtg	ctggaggaga	acagggtagg	14580
atgggtgtggg	gctttttgtag	actgtctaga	agcaaagaag	gtctgcagtg	gaggtggtgt	14640
ctgaggtgaa	tctcgaagg	gaataggagt	tgaacgttag	caggcagagg	gtggattgca	14700
ggagagcagc	ggcctgggca	ggtgccacgc	gtggcccatc	aggggtgcttc	atgcatggct	14760
gtgtgcttgc	catccttct	gcctgcctac	ccctgctgc	ttcgcttcat	gggggcgttt	14820
gagcttgggc	ccacctgcct	gcctcgcttg	tgggcagagg	acccaagctg	tgtgagttgt	14880
cctgtcccgg	ggagcagctg	aactgggtccg	gggtctcgaa	ctgtggggct	caaaaggact	14940
ccggggctcat	ttcaactggg	ctgtgccgat	tcctgggggc	tgttnggaan	gtaaaggcct	15000
aaaggggctc	cctggttang	gccctcaant	ttaanaacct	ggggccgggg	cccqgaattg	15060

cccccaantt tgtttcaacn ccccttggcc ttnggcnggg gcaaatttcc anggggaacc 15120  
aatggntttc ccccaaaaan ggggcnttt taaccnttt ccaaanttg ggcctaaaa 15180  
aagggtggan ttcctgaang gg 15202

<210> 22  
<211> 1070  
<212> PRT  
<213> Homo sapiens

<400> 22

Val	Cys	Leu	Asp	Asp	Thr	Val	Thr	Thr	Glu	Lys	Glu	Leu	Asp	Thr	Val
1				5					10					15	
Glu	Val	Leu	Lys	Ala	Ile	Gln	Lys	Ala	Lys	Glu	Val	Lys	Ser	Lys	Leu
			20					25					30		
Ser	Asn	Pro	Glu	Lys	Lys	Gly	Gly	Glu	Asp	Ser	Arg	Leu	Ser	Ala	Ala
		35					40					45			
Pro	Cys	Ile	Arg	Pro	Ser	Ser	Ser	Pro	Pro	Thr	Val	Ala	Pro	Ala	Ser
	50					55					60				
Ala	Ser	Leu	Pro	Gln	Pro	Ile	Leu	Ser	Asn	Gln	Gly	Ile	Met	Phe	Val
65					70					75					80
Gln	Glu	Glu	Ala	Leu	Ala	Ser	Ser	Leu	Ser	Ser	Thr	Asp	Ser	Leu	Thr
				85					90					95	
Pro	Glu	His	Gln	Pro	Ile	Ala	Gln	Gly	Cys	Ser	Asp	Ser	Leu	Glu	Ser
			100					105					110		
Ile	Pro	Ala	Gly	Gln	Ala	Ala	Ser	Asp	Asp	Leu	Arg	Asp	Val	Pro	Gly
		115					120					125			
Ala	Val	Gly	Gly	Ala	Ser	Pro	Glu	His	Ala	Glu	Pro	Glu	Val	Gln	Val
	130					135					140				
Val	Pro	Gly	Ser	Gly	Gln	Ile	Ile	Phe	Leu	Pro	Phe	Thr	Cys	Ile	Gly
145					150					155					160
Tyr	Thr	Ala	Thr	Asn	Gln	Asp	Phe	Ile	Gln	Arg	Leu	Ser	Thr	Leu	Ile
				165					170					175	
Arg	Gln	Ala	Ile	Glu	Arg	Gln	Leu	Pro	Ala	Trp	Ile	Glu	Ala	Ala	Asn
			180					185					190		
Gln	Arg	Glu	Glu	Gly	Gln	Gly	Glu	Gln	Gly	Glu	Glu	Glu	Asp	Glu	Glu
		195					200					205			
Glu	Glu	Glu	Glu	Glu	Asp	Val	Ala	Glu	Asn	Arg	Tyr	Phe	Glu	Met	Gly
	210					215					220				
Pro	Pro	Asp	Val	Glu	Glu	Glu	Glu	Gly	Gly	Gln	Gly	Glu	Glu	Glu	
225					230				235						240

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted March 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Glu	Glu	Glu	Glu	Glu	Asp	Glu	Glu	Ala	Glu	Glu	Glu	Arg	Leu	Ala	Leu	
				245					250					255		
Glu	Trp	Ala	Leu	Gly	Ala	Asp	Glu	Asp	Phe	Leu	Leu	Glu	His	Ile	Arg	
			260					265					270			
Ile	Leu	Lys	Val	Leu	Trp	Cys	Phe	Leu	Ile	His	Val	Gln	Gly	Ser	Ile	
		275					280					285				
Arg	Gln	Phe	Ala	Ala	Cys	Leu	Val	Leu	Thr	Asp	Phe	Gly	Ile	Ala	Val	
	290					295					300					
Phe	Glu	Ile	Pro	His	Gln	Glu	Ser	Arg	Gly	Ser	Ser	Gln	His	Ile	Leu	
305					310					315					320	
Ser	Ser	Leu	Arg	Phe	Val	Phe	Cys	Phe	Pro	His	Gly	Asp	Leu	Thr	Glu	
				325					330					335		
Phe	Gly	Phe	Leu	Met	Pro	Glu	Leu	Cys	Leu	Val	Leu	Lys	Val	Arg	His	
			340					345					350			
Ser	Glu	Asn	Thr	Leu	Phe	Ile	Ile	Ser	Asp	Ala	Ala	Asn	Leu	His	Glu	
		355					360					365				
Phe	His	Ala	Asp	Leu	Arg	Ser	Cys	Phe	Ala	Pro	Gln	His	Met	Ala	Met	
	370					375					380					
Leu	Cys	Ser	Pro	Ile	Leu	Tyr	Gly	Ser	His	Thr	Ser	Leu	Gln	Glu	Phe	
385					390					395					400	
Leu	Arg	Gln	Leu	Leu	Thr	Phe	Tyr	Lys	Val	Ala	Gly	Gly	Cys	Gln	Glu	
				405					410					415		
Arg	Ser	Gln	Gly	Cys	Phe	Pro	Val	Tyr	Leu	Val	Tyr	Ser	Asp	Lys	Arg	
			420					425					430			
Met	Val	Gln	Thr	Ala	Ala	Gly	Asp	Tyr	Ser	Gly	Asn	Ile	Glu	Trp	Ala	
		435					440					445				
Ser	Cys	Thr	Leu	Cys	Ser	Ala	Val	Arg	Arg	Ser	Cys	Cys	Ala	Pro	Ser	
	450					455					460					
Glu	Ala	Val	Lys	Ser	Ala	Ala	Ile	Pro	Tyr	Trp	Leu	Leu	Leu	Thr	Pro	
465					470					475					480	
Gln	His	Leu	Asn	Val	Ile	Lys	Ala	Asp	Phe	Asn	Pro	Met	Pro	Asn	Arg	
			485						490					495		
Gly	Thr	His	Asn	Cys	Arg	Asn	Arg	Asn	Ser	Phe	Lys	Leu	Ser	Arg	Val	
			500					505					510			
Pro	Leu	Ser	Thr	Val	Leu	Leu	Asp	Pro	Thr	Arg	Ser	Cys	Thr	Gln	Pro	
		515					520					525				
Arg	Gly	Ala	Phe	Ala	Asp	Gly	His	Val	Leu	Glu	Leu	Leu	Val	Gly	Tyr	
	530					535					540					
Arg	Phe	Val	Thr	Ala	Ile	Phe	Val	Leu	Pro	His	Glu	Lys	Phe	His	Phe	
545					550					555					560	
Leu	Arg	Val	Tyr	Asn	Gln	Leu	Arg	Ala	Ser	Leu	Gln	Asp	Leu	Lys	Thr	
				565					570					575		

Val	Val	Ile	Ala	Lys	Thr	Pro	Gly	Thr	Gly	Gly	Ser	Pro	Gln	Gly	Ser		
			580					585					590				
Phe	Ala	Asp	Gly	Gln	Pro	Ala	Glu	Arg	Arg	Ala	Ser	Asn	Asp	Gln	Arg		
		595					600					605					
Pro	Gln	Glu	Val	Pro	Ala	Glu	Ala	Leu	Ala	Pro	Ala	Pro	Val	Glu	Val		
	610					615					620						
Pro	Ala	Pro	Ala	Pro	Ala	Ala	Ala	Ser	Ala	Ser	Gly	Pro	Ala	Lys	Thr		
625					630					635					640		
Pro	Ala	Pro	Ala	Glu	Ala	Ser	Thr	Ser	Ala	Leu	Val	Pro	Glu	Glu	Thr		
				645					650					655			
Pro	Val	Glu	Ala	Pro	Ala	Pro	Pro	Pro	Ala	Glu	Ala	Pro	Ala	Gln	Tyr		
			660					665						670			
Pro	Ser	Glu	His	Leu	Ile	Gln	Ala	Thr	Ser	Glu	Glu	Asn	Gln	Ile	Pro		
		675					680					685					
Ser	His	Leu	Pro	Ala	Cys	Pro	Ser	Leu	Arg	His	Val	Ala	Ser	Leu	Arg		
	690					695					700						
Gly	Ser	Ala	Ile	Ile	Glu	Leu	Phe	His	Ser	Ser	Ile	Ala	Glu	Val	Glu		
705					710					715					720		
Asn	Glu	Glu	Leu	Arg	His	Leu	Met	Trp	Ser	Ser	Val	Val	Phe	Tyr	Gln		
				725					730					735			
Thr	Pro	Gly	Leu	Glu	Val	Thr	Ala	Cys	Val	Leu	Leu	Ser	Thr	Lys	Ala		
			740					745						750			
Val	Tyr	Phe	Val	Leu	His	Asp	Gly	Leu	Arg	Arg	Tyr	Phe	Ser	Glu	Pro		
		755					760					765					
Leu	Gln	Asp	Phe	Trp	His	Gln	Lys	Asn	Thr	Asp	Tyr	Asn	Asn	Ser	Pro		
	770					775					780						
Phe	His	Ile	Ser	Gln	Cys	Phe	Val	Leu	Lys	Leu	Ser	Asp	Leu	Gln	Ser		
785					790					795					800		
Val	Asn	Val	Gly	Leu	Phe	Asp	Gln	His	Phe	Arg	Leu	Thr	Gly	Ser	Thr		
				805					810					815			
Pro	Met	Gln	Val	Val	Thr	Cys	Leu	Thr	Arg	Asp	Ser	Tyr	Leu	Thr	His		
			820					825					830				
Cys	Phe	Leu	Gln	His	Leu	Met	Val	Val	Leu	Ser	Ser	Leu	Glu	Arg	Thr		
		835					840					845					
Pro	Ser	Pro	Glu	Pro	Val	Asp	Lys	Asp	Phe	Tyr	Ser	Glu	Phe	Gly	Asn		
	850					855					860						
Lys	Thr	Thr	Gly	Lys	Met	Glu	Asn	Tyr	Glu	Leu	Ile	His	Ser	Ser	Arg		
865					870					875					880		
Val	Lys	Phe	Thr	Tyr	Pro	Ser	Glu	Glu	Glu	Ile	Gly	Asp	Leu	Thr	Phe		
				885					890					895			
Thr	Val	Ala	Gln	Lys	Met	Ala	Glu	Pro	Glu	Lys	Ala	Pro	Ala	Leu	Ser		
			900					905					910				

